CDF/PUB/TOP/CDFR/11076

First Search for *s*-channel Electroweak Single Top Quark Production in the Missing Energy Plus Jets Topology and Combination with the Same Search in the Lepton Plus Jets Topology using the Full CDF II Data Set

The CDF Collaboration[1]

[1] *URL http://www-cdf.fnal.gov*

(Dated: January 30, 2014)

Abstract

The first search for electroweak single top quark production from the exchange of an s-channel virtual W boson using a data sample composed of events with an imbalance in the total transverse momentum, b-tagged jets, and no identified leptons ($\not\!\!E_T b\bar{b}$) is presented. The full CDF II data set, corresponding to an integrated luminosity of 9.45 fb^{-1} from Fermilab Tevatron proton-antiproton collisions at a center of mass energy of 1.96 TeV, is used and an excess of events with respect to estimated standard-model backgrounds is observed. Assuming that this excess is due to electroweak production of top quarks of mass 172.5 GeV/c^2 in the s-channel, a cross section of $1.12^{+0.61}_{-0.57}$ (stat+syst) pb with a significance of 1.9 standard deviations is measured. This measurement is combined with a previously reported CDF result obtained from events with an imbalance in total transverse momentum, b-tagged jets, and exactly one identified lepton ($\ell\nu b\bar{b}$), yielding an s-channel electroweak single top quark cross section of $1.36^{+0.37}_{-0.32}$ (stat+syst) pb with a significance of 4.2 standard deviations.

The top quark was discovered at Fermilab in 1995 [1, 2] in top-antitop-quark pair production. This process is mediated by the strong interaction and results in the largest contribution to the top-quark-production cross section in hadron collisions. The top quark can also be produced singly via the electroweak interaction involving the Wtb vertex. The study of electroweak single top quark production is particularly interesting because of the direct dependence of the cross section on the magnitude of the Wtb coupling. Furthermore, electroweak single top quark production from the exchange of an s-channel virtual W boson is of special interest since possible deviations from the standard model (SM) expectation could indicate evidence for new physics such as the existence of higher-mass bosons of the W family (W') or a charged Higgs boson [3]. Examples of SM electroweak single top production

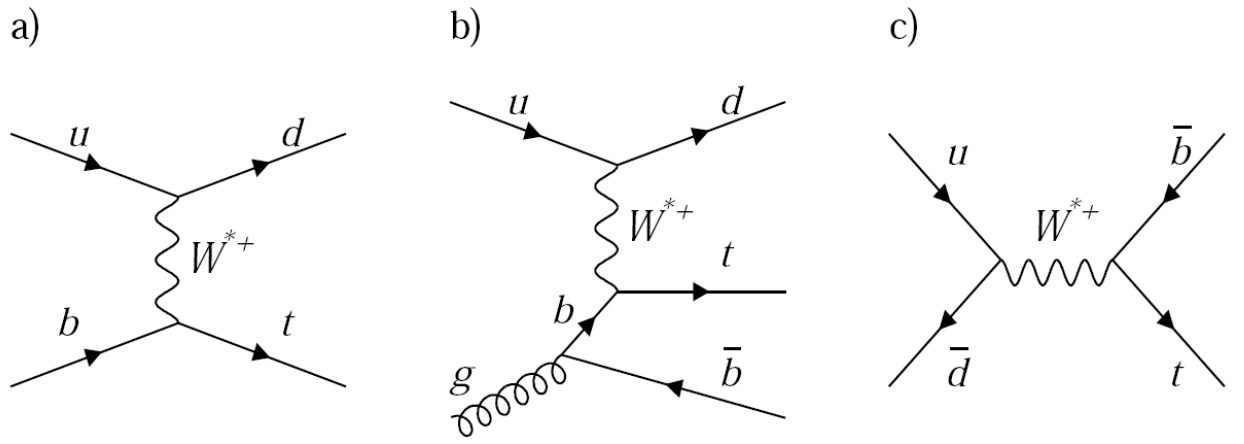

FIG. 1. Feynman diagrams for electroweak single top quark production: a) leading-order t-channel, b) next-to-leading-order t-channel production, and c) leading-order s-channel production.

diagrams dominating at the Tevatron are shown in Figure 1.

Electroweak single top quark production was observed at the Tevatron in 2009 [4–6] in the combined t- and s-channels. However, s-channel production has yet to be observed independently. While both the associated production of an on-shell W boson and a top quark (Wt-channel, negligible at the Tevatron) [7] and electroweak single top quark production through the t-channel exchange of a W boson [8, 9] were established by the ATLAS and CMS Collaborations, the s-channel process has an unfavorable production rate compared to the background rates at the LHC. The D0 Collaboration reported the first evidence of s-channel electroweak single top quark production [10]. More recently CDF also obtained 3.8

standard deviations evidence using a sample of events containing one isolated lepton (e or μ), large missing transverse energy [11] ($\not\!\!E_T$), and two jets, at least one of which is identified as likely to have originated from a bottom quark (b-tagged), referred as the $\ell\nu b\bar{b}$ sample [12]. In this letter, the same search, performed for the first time using an independent sample of events with large $\not\!\!E_T$, two or three jets, with one or more b-tagged jets, and no detected e or μ lepton candidates ($\not\!\!E_T b\bar{b}$ sample) is reported. Most of the techniques developed for the low-mass Higgs boson search in the same data sample are exploited, including the HOBIT [13] b-tagger. By combining the results of the two searches, the best possible sensitivity to s-channel single top production from the 9.45 fb^{-1} of integrated luminosity of the full CDF II data set is obtained.

In the $\not\!\!E_T b\bar{b}$ analysis, events are accepted by the online event selection (trigger) by requiring $\not\!\!E_T > 45$ GeV or alternatively $\not\!\!E_T > 35$ GeV and two or more jets with transverse energy [11] $E_T > 15$ GeV. Offline, events containing identified electrons or muons are excluded and $\not\!\!E_T > 35$ GeV is required, after correcting measured jet energies for instrumental effects [14]. Events with two or three high-E_T jets are selected and the two jets with the largest transverse energies, $E_T^{j_1}$ and $E_T^{j_2}$, are required to satisfy $25 < E_T^{j_1} < 200$ GeV and $20 < E_T^{j_2} < 120$ GeV, where the jet energies are determined from calorimeter deposits corrected for track momentum measurements [15]. This sample contains some events consistent with electroweak single top quark production where the tau lepton from the $t \to Wb \to \tau\nu b$ decay is reconstructed as a jet in the calorimeters. To increase the acceptance for events with an unidentified τ lepton, events in which the third-most energetic jet satisfies $15 < E_T^{j_3} < 100$ GeV are accepted. Because of the large rate of inclusive quantum chromodynamics (QCD) multijet (MJ) production, events with four reconstructed jets where each jet has transverse energy in excess of 15 GeV and pseudorapidity [11] $|\eta| < 2.4$ are rejected. To ensure that the two leading-E_T jets are within the silicon-detector acceptance, they are required to satisfy $|\eta| < 2$, with at least one of them satisfying $|\eta| < 0.9$.

The MJ background events most often contain $\not\!\!E_T$ generated through jet energy mismeasurements. Neutrinos produced in semileptonic b-hadron decays can also contribute to the $\not\!\!E_T$ observed within these events. In both cases, the $\not\!\!E_T$ is typically aligned with $\vec{E}_{\rm T}^{j_2}$ and events are rejected by requiring $\Delta\varphi(\vec{\not\!\!E}_T\ , \vec{E}_{\rm T}^{j_{2,3}}) \geq 0.4$. The remaining MJ background has a large contribution of events with jets from fragmenting light-flavoured (u, d, s) quarks or

gluons which can be further reduced by requiring b-tagged jets. Charm quarks, which share some features associated with b quarks, are not explicitly identified. Events are assigned to three independent subsamples depending on the HOBIT output of the two leading jets. Jets with HOBIT values larger than 0.98 are defined as tightly tagged (T-jet), whereas jets with HOBIT outputs between 0.72 and 0.98 are defined as loosely tagged (L-jet). TT events are defined as those in which both jets are tightly tagged, TL events as those in which one jet is tightly tagged and the other is loosely tagged, and 1T events as those in which only one jet is tightly tagged while the other is untagged. Events with two or three jets are analyzed separately, leading to six event subsamples with differing signal to background ratios. This strategy enhances sensitivity and helps separate s-channel electroweak single top quark production, enhanced in the double tagging categories, from the t-channel production, enhanced in the single tagging categories.

In order to extract the s-channel electroweak single top quark signal from the more dominant background contributions, the rates and kinematic distributions of events associated with each process need to be accurately modeled. The kinematic distributions of events associated with top-quark, V+jets (where V stands for W and Z bosons), $W + c$, diboson (VV) and associated Higgs and W or Z boson (VH) production are modeled using computer simulations. The ALPGEN generator [16] is used to model V+jets, $W+c$, and VH production. The POWHEG [17] generator is used to model t-and s-channel electroweak single top quark production, while PYTHIA [18] is used to model top-quark pair and VV production. Parton showering is simulated in all cases using PYTHIA. Event modeling includes simulation of the detector response using GEANT [19]. The simulated events are reconstructed and analyzed in the same way as the experimental data. Normalizations of the event contributions from t-channel electroweak single top quark, VV, VH, and $t\bar{t}$ pair production are taken from theoretical cross section predictions [20–23], while normalization for $W+c$ production is taken from the measured cross section [24]. For V+jets production the heavy-flavor contribution is normalized based on the number of b-tagged events observed in an orthogonal data control sample. Contributions of electroweak mistags, V+jets, and VV events containing at least one incorrectly b-tagged, light-flavored jet are determined by applying per-event mistag probabilities obtained from a generic event sample containing light-flavored jets to simulated events [25]. The MJ background [26] remaining after application of the full selection criteria

is modeled by applying a tag-rate matrix derived from a MJ-dominated data sample to pre-tagged events that otherwise satisfy the signal sample selection criteria.

To separate the s-channel electroweak single top quark signal from the backgrounds, a staged multivariate neural network (NN) technique is employed. A first NN, $\mathrm{NN_{QCD}}$, is trained to discriminate MJ events from signal events. Events that satisfy a minimal requirement on the $\mathrm{NN_{QCD}}$ output variable are further analyzed by a function, $\mathrm{NN_{sig}}$, derived from the outputs of two additional NNs, $\mathrm{NN}_{V\mathrm{jets}}$ and $\mathrm{NN}_{\mathrm{t\bar{t}}}$, designed respectively to separate the signal from V+jets (and the remaining MJ events) and $t\bar{t}$ backgrounds.

Since the kinematic properites associated with the presence of a W boson in the s-channel electroweak single top quark and W+jets production processes are very similar, in contrast with those of events absent a W boson originating from MJ production, the $\mathrm{NN_{QCD}}$ discriminant is trained using QCD multijet events for the background sample and W+jets events for the signal sample. The discriminant is trained separately for the two-jet and three-jet samples using kinematic, angular, and event-shape quantities for the input variables.

The two additional NNs, $\mathrm{NN}_{V\mathrm{jets}}$ and $\mathrm{NN}_{\mathrm{t\bar{t}}}$, are trained for events that satisfy the minimum requirement on the $\mathrm{NN_{QCD}}$ output variable. The first, $\mathrm{NN}_{V\mathrm{jets}}$, is trained to separate the s-channel electroweak single top quark signal from V+jets and the remaining MJ backgrounds using simulated s-channel electroweak single top quark events for the signal sample and pre-tagged data events that satisfy the requirement on $\mathrm{NN_{QCD}}$, reweighted by the tag rate parameterization, for the background sample. The $\mathrm{NN_{QCD}}$ requirement changes the pre-tag data composition, enhancing the V+jets contribution and selecting MJ events with properties closer to those expected for V+jets events. The background model obtained by re-weighting these events via the tag-rate probability function accounts for both the V+jets and MJ event contributions, allowing for more straight-forward training of the $\mathrm{NN}_{V\mathrm{jets}}$. The second, $\mathrm{NN}_{\mathrm{t\bar{t}}}$, is trained to separate s-channel electroweak single top quark from $t\bar{t}$ production using simulated s-channel electroweak single top quark for the signal and simulated $t\bar{t}$ for the background. The final discriminant, $\mathrm{NN_{sig}}$, is defined as the quadrature-weighted sum of the $\mathrm{NN}_{V\mathrm{jets}}$ and $\mathrm{NN}_{\mathrm{t\bar{t}}}$ output variables. Figure 2 shows the predicted and observed shapes of $\mathrm{NN_{sig}}$ output variable for each of the six event subsamples used in this search.

The modeling of SM backgrounds is tested in several control samples. A first (EWK) control sample is defined containing events with at least one charged lepton that otherwise

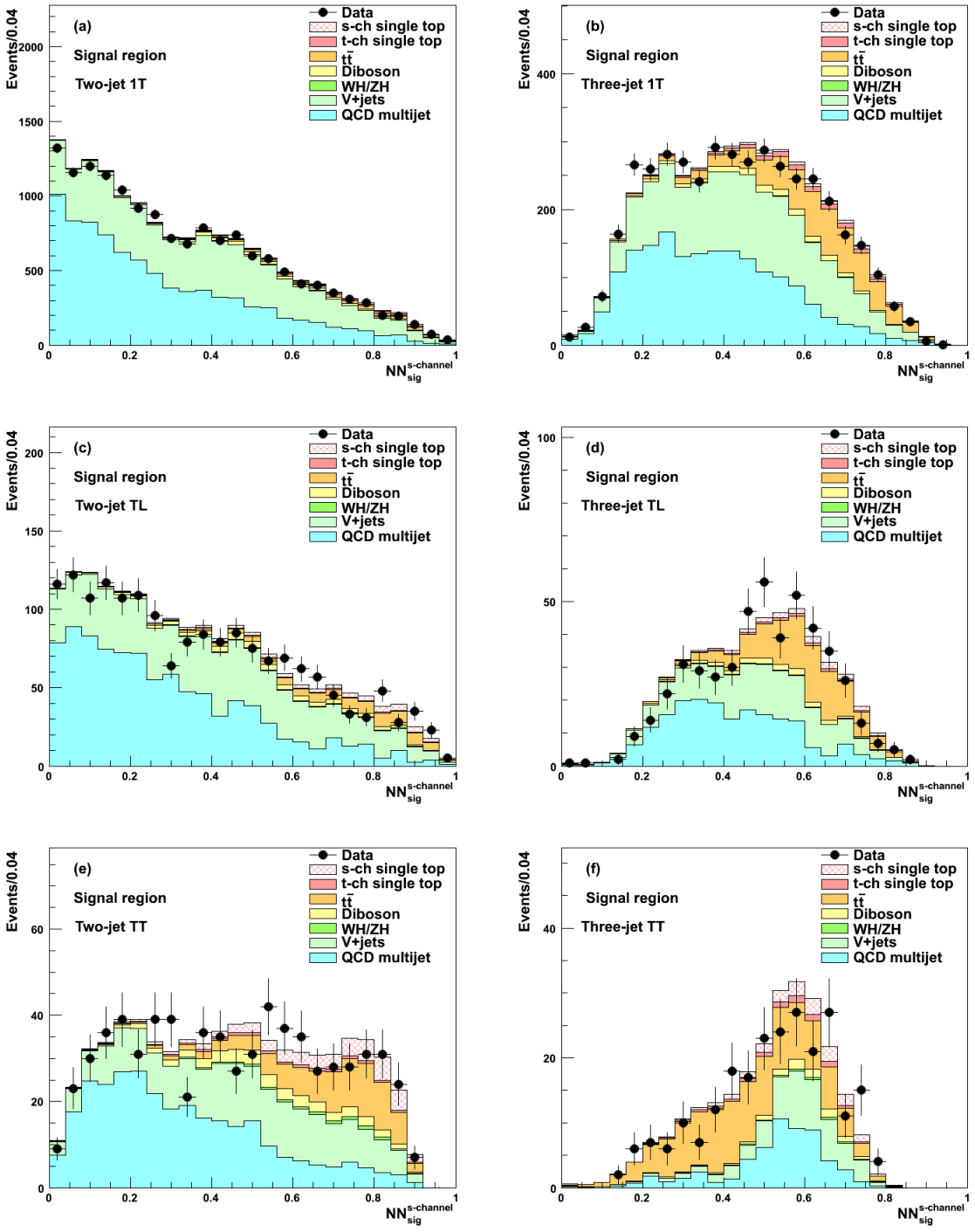

FIG. 2. Predicted and observed final discriminant distributions in the signal region, for 1T two-jet (a), 1T three-jet (b), TL two-jet (c), TL three-jet (d), TT two-jet (e) and TT three-jet (f) subsamples.

TABLE I. Number of predicted and observed events in the two-jet signal region in the subsample with exactly one tight HOBIT tagged jet (1T), one tight and one loose HOBIT tagged jet (TL) and two tight HOBIT tagged jets (TT). The uncertainties in the predicted numbers of events are due to the theoretical-cross-section uncertainties and to the uncertainties on signal and background modeling. Both the uncertainties and the central values are those given by the fit to the data with theory constraints.

Category	1T	TL	TT
MJ	8238.0 ± 95.5	942.0 ± 43.1	292.8 ± 14.2
V+jets	6174.7 ± 682.6	631.9 ± 75.1	193.8 ± 21.7
t-ch single top	154.1 ± 28.8	17.5 ± 3.4	10.6 ± 2.0
$t\bar{t}$	261.4 ± 23.9	98.7 ± 9.9	102.5 ± 9.4
Diboson	297.4 ± 40.4	50.0 ± 6.9	37.7 ± 4.9
Signal	85.3 ± 16.6	42.8 ± 8.5	46.1 ± 9.0
Total prediction	15210.9 ± 691.6	1782.9 ± 87.9	683.5 ± 29.5
Observed	15312	1743	686

satisfy the selection criteria. This sample is independent from the signal sample and is sensitive primarily to top-quark pair, V+jets, and, to a lesser extent, VV production. A second (QCD) control sample contains events that do not meet the minimal requirement on the $\mathrm{NN_{QCD}}$ output variable but otherwise satisfy the search sample selection criteria. This event sample, dominated by MJ production, is used to validate the data-driven MJ background model and obtain scale factors, ranging from 0.7 to 0.9, for normalizing modeled contributions to the TT, TL, and 1T event subsamples. Comparisons of modeled and observed distributions for multiple kinematic variables, including those used as inputs to the $\mathrm{NN_{QCD}}$, $\mathrm{NN}_{V\mathrm{jets}}$, and $\mathrm{NN}_{\mathrm{t\bar{t}}}$, are used to validate the accuracy of the model.

To measure the signal contribution, the sum of modeled contributions is fitted to the observed data as a function of the final discriminant variable, $\mathrm{NN_{sig}}$, accounting for statistical and systematic uncertainties. The dominant systematic uncertainties arise from the normalization of the V-plus-heavy-flavor background contributions (30%), differences in b-tagging efficiencies between data and simulation (8–16%) and mistag rates (20% –30%). [13]. Other

TABLE II. Number of predicted and observed events in the three-jet signal region in the subsample with exactly one tight HOBIT tagged jet (1T), one tight and one loose HOBIT tagged jet (TL) and two tight HOBIT tagged jets (TT). The uncertainties in the predicted numbers of events are due to the theoretical-cross-section uncertainties and to the uncertainties on signal and background modeling. Both the uncertainties and the central values are those given by the fit to the data with theory constraints.

Category	1T	TL	TT
MJ	1733.6 ± 34.2	206.7 ± 13.2	58.2 ± 5.3
V+jets	1550.8 ± 171.8	90.2 ± 10.9	13.4 ± 2.1
t-ch single top	79.1 ± 14.8	10.8 ± 2.1	7.4 ± 1.4
$t\bar{t}$	636.1 ± 58.0	148.5 ± 14.9	124.3 ± 11.4
Diboson	111.3 ± 15.2	14.6 ± 2.0	8.7 ± 1.2
Signal	44.9 ± 8.8	16.2 ± 3.2	16.5 ± 3.2
Total prediction	4155.8 ± 186.0	487.0 ± 23.1	228.5 ± 13.3
Observed	4198	490	237

uncertainties are on the $t\bar{t}$ (3.5%), t-channel electroweak single top quark (6.2%), VV (6%), VH (5%), and $W + c$ (23%) cross sections [20–24], normalizations of the QCD multijet background (3–7%), luminosity measurement (6%) [27], jet-energy scale (1–6%) [14], trigger efficiency (1–3%), parton distribution functions (2%), and lepton vetoes (2%). The shapes obtained by varying the tag-rate probabilities by one standard deviation from their central values are applied as uncertainties on shapes of the $\mathrm{NN_{sig}}$ output distribution for the MJ background. Changes in the shape of the $\mathrm{NN_{sig}}$ distribution originating from jet energy scale uncertainties are also incorporated for processes modeled via the simulation.

A likelihood fit to the binned $\mathrm{NN_{sig}}$ distribution is used to extract an s-channel electroweak single top quark signal in the presence of SM backgrounds. The likelihood is the product of Poisson probabilities over the bins of the final discriminant distribution. The mean number of expected events in each bin includes contributions from each background source and s-channel electroweak single top quark production, assuming a top quark mass of 172.5 GeV/c^2. To extract the most likely s-channel electroweak single top quark contribution, a

Bayesian method [28] is employed. A uniform prior probability in the non-negative range for the s-channel electroweak single top quark production cross section times branching fraction and truncated Gaussian priors for the uncertainties on the acceptance and shape of each process are incorporated within the fit. Results from each of the six search subsamples are combined by taking the product of their likelihoods and simultaneously varying the correlated uncertainties.

The measured s-channel electroweak single top quark cross section in the $\not\!\!E_T b\bar{b}$ sample is $1.12^{+0.61}_{-0.57}$ pb, where the probability of observing a signal cross section larger than the observed one assuming that the true cross section is zero as determined from background-only pseudoexperiments (p-value) is 3.1×10^{-2}, corresponding to a significance of 1.9 standard deviations.

This result is combined with the result of a CDF search for s-channel electroweak single top quark production in the $\ell\nu b\bar{b}$ sample, as described in Ref. [12]. In this search, candidate events were selected by requiring exactly one reconstructed charged lepton (e or μ) in the final state. Hence, no such events are included in the $\not\!\!E_T b\bar{b}$ analysis described above. Four independent tagging categories, according to the score of the HOBIT tagger on the two leading jets (tight-tight TT, tight-loose TL, single-tight 1T, and loose-loose LL), were analyzed separately. Events were also divided into three independent categories based on different lepton reconstruction algorithms. To further discriminate the signal from all other backgrounds, NNs were employed. These NNs were optimized separately for each tagging and lepton category. Correlated systematic errors were treated as described above for the $\not\!\!E_T b\bar{b}$ search. Finally, a Bayesian binned-likelihood technique was applied to the final NN output to extract the s-channel electroweak single top quark cross section. The significance of the result from the $\ell\nu b\bar{b}$ channel was 3.8 standard deviations, and the measured cross section was $1.41^{+0.44}_{-0.42}$ (stat+syst) pb, assuming a top-quark mass of 172.5 GeV/c^2.

The two analyses are combined by taking the product of their likelihoods and simultaneously varying the correlated uncertainties, following the same procedure explained above. The uncertainties associated with the theoretical cross sections of the $t\bar{t}$, t-channel electroweak single top quark, VV, and VH production processes; the luminosity; the b-tagging efficiency; and the mistag rate are taken as fully correlated between the two searches. The combined measurement results in an s-channel electroweak single top quark production cross section of $1.36^{+0.37}_{-0.32}$ pb, consistent with the SM cross section of 1.05 ± 0.05 pb [21]. The combined

p-value is 1.6×10^{-5}, which corresponds to a signal significance of 4.2 standard deviations. The median expected significance assuming that a signal is present at the SM rate is 3.4 standard deviations. This result indicates that the signal observed in the $\not\!\!E_T b\bar{b}$ sample further strengthens the evidence for s-channel electroweak single top quark production reported in Ref.[12], improving the uncertainty on the measured cross section by more than 10%.

We thank the Fermilab staff and the technical staffs of the participating institutions for their vital contributions. This work was supported by the U.S. Department of Energy and National Science Foundation; the Italian Istituto Nazionale di Fisica Nucleare; the Ministry of Education, Culture, Sports, Science and Technology of Japan; the Natural Sciences and Engineering Research Council of Canada; the National Science Council of the Republic of China; the Swiss National Science Foundation; the A.P. Sloan Foundation; the Bundesministerium für Bildung und Forschung, Germany; the Korean World Class University Program, the National Research Foundation of Korea; the Science and Technology Facilities Council and the Royal Society, UK; the Russian Foundation for Basic Research; the Ministerio de Ciencia e Innovación, and Programa Consolider-Ingenio 2010, Spain; the Slovak R&D Agency; the Academy of Finland; the Australian Research Council (ARC); and the EU community Marie Curie Fellowship contract 302103.

[1] F. Abe *et al.* (CDF Collaboration), Phys. Rev. Lett. **74**, 2626 (1995).

[2] S. Abachi *et al.* (D0 Collaboration), Phys. Rev. Lett. **74**, 26322637 (1995).

[3] T. M. P. Tait and C. P. Yuan, Phys. Rev. D **63**, 014018 (2000).

[4] CDF and D0 Collaboration, Tevatron Electroweak Working Group, arXiv:0908.2171.

[5] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **103**, 092002 (2009).

[6] V. M. Abazov *et al.* (D0 Collaboration), Phys. Rev. Lett. **103**, 092001 (2009).

[7] S. Chatrchyan *et al.* (CMS Collaboration), Phys. Rev. Lett. **110**, 022003 (2013).

[8] ATLAS Collaboration, ATLAS-CONF-2011-101 (2011).

[9] CMS Collaboration, CMS-TOP-10-008 (2011).

[10] V. M. Abazov *et al.* (D0 Collaboration), Phys. Lett. B **726** 656 (2013).

[11] CDF uses a cylindrical coordinate system with the z axis along the proton beam axis. The pseudorapidity is $\eta = -\ln(\tan(\frac{\theta}{2}))$, where θ is the polar angle, and φ is the azimuthal angle,

while $p_{\rm T} = p\sin\theta$ and $E_{\rm T} = E\sin\theta$. The $\not\!\!E_T$ is defined as the magnitude of $\vec{\not\!\!E}_T = -\sum_i E_T^i \hat{n}_i$, where $\hat{n}_i$ is a unit vector perpendicular to the beam axis and pointing at the ith calorimeter tower, and E_T^i is the transverse energy therein.

[12] T. Aaltonen *et al.* (CDF Collaboration), To be Published in Phys. Rev. Lett.

[13] J. Freeman, T. Junk, M. Kirby, Y. Oksuzian, T. J. Phillips, F. D. Snider, M. Trovato, J. Vizan, and W. M. Yao, Nucl. Instrum. Methods Phys. Res., Sect. A **697**, 64 (2013).

[14] A. Bhatti *et al.*, Nucl. Instrum. Methods Phys. Res., Sect. A **566**, 375 (2006).

[15] C. Adloff *et al.* (H1 Collaboration), Z. Phys. C **74**, 221 (1997).

[16] M.L. Mangano, M. Moretti, F. Piccinini, R. Pittau, and A.D. Polosa, J. High Energy Phys. **0307**, 001 (2003).

[17] S. Alioli *et al.*, J. High Energy Phys. 1006 (2010) 043 .

[18] T. Sjostrand, S. Mrenna, and P. Skands, J. High Energy Phys. 05 (2006) 026. We use PYTHIA version 6.216 to generate the Higgs boson signals.

[19] GEANT, Detector description and simulation tool, CERN Program Library Long Writeup W5013 (1993).

[20] P. Baernreuther, M. Czakon and A. Mitov, Phys. Rev. Lett. **109**, 132001 (2012).

[21] N. Kidonakis, Phys. Rev. D **81**, 054028 (2010).

[22] J. M. Campbell and R. K. Ellis, Phys. Rev. D **60**, 113006 (1999).

[23] J. Baglio and A. Djouadi, J. High Energy Phys. 10 (2010) 064; O. Brien, R. V. Harlander, M. Weisemann, and T. Zirke, Eur. Phys. J. C **72**, 1868 (2012).

[24] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. Lett. **110**, 071801 (2013).

[25] D. Acosta *et al.* (CDF Collaboration), Phys. Rev. D **71**, 052003 (2005); A. Abulencia *et al.* (CDF Collaboration), Phys. Rev. D **74**, 072006 (2006).

[26] T. Aaltonen *et al.* (CDF Collaboration), Phys. Rev. D **87**, 052008 (2013).

[27] S. Klimenko, J. Konigsberg, and T. M. Liss, Report No. FERMILAB-FN-0741, 2003.

[28] *Statistics*, in K. Nakamura *et al.* (Particle Data Group), J. Phys. G **37** 075021 (2010).